THE MOLECULAR PARADIGM IN BIOLOGY
A Synopsis

I. THE GENETIC POINT OF VIEW

Origins.

The beginnings of modern biology can be traced to the new spirit of curiosity that emerged in Europe during the Renaissance. The transformation in art preceded that in science. We see this in the paintings (e.g. Giorgione, Bellini) where scenes from nature as well as religious themes are depicted with a fresh eye for realistic detail. Secular and religious are symbolically united in Leonardo's "Annunciation": set in the out-of-doors and uniting exquisite observation of natural detail with a sense of energy and religious mystery. Order in nature and the transformations of energy fascinated scientists of the 16th and 17th centuries. The drama of Galileo's clash with the Inquisition should not divert us from seeing the tension within the scientist himself between religious faith in divine order and a new-found empiricism in investigating natural phenomena. Newton typifies this tension. The properties of matter and its interaction with energy fired his scientific imagination and his sense of divine mission. He was a precise observer according to the Baconian precept and he coupled this with the formulation of quantitative laws amenable to mathematical treatment. Thus his physics became the model for how scientific explanation worked, at a time when physiology and microbiology (Harvey, Leeuwenhoek) were still largely observational.

The "interior mold".

Post-Renaissance advances in physical sciences led to a number of complex but speculative theories of organisms as automata, beginning with Descartes. In Diderot's fantasy d'Alembert, waking from a dream, asks how he remains himself despite the turnover of matter in his organs. "Piecemealness" is the answer, the organism being compared to a monastery gradually endoctrinating new recruits. A similar concept of body turnover was advanced by La Mettrie (L'homme Machine), who held a distinctly modern view of phsyiological (and psychological) phenomena arising from particular configurations of matter. The then dominant "fiber" theory of bodily structure also has a modern flavor, despite the faulty view of histology to which it was attached.

The stability and self-generative capacity of biological systems suggested early in the development of the science that these characteristics depended on an internal ordering principle or plan. In one sense, this

notion can be traced to philosophical and theological systems in which objects as perceived reflect ideal forms or divine essences. From the modern point of view, we are accustomed to expecting any complex machine to have a blueprint and to be reducible to physical principles. [Theories of self-duplicating automata have been developed (von Neumann, Wiener) but have not been especially useful in biology.] It appears reasonable that if a structure is to be in dynamic interaction with an environment, yet stable over time (or changing in a pre-determined manner), there must be a process of checking what is happening against a master plan of some sort.

Something akin to the modern concept of a genetic template is to be found in the speculative writings of Buffon. He postulated a hierarchy of "interior molds" to explain the incorporation of nutrients into the fabric of the organism, and reproduction of new individuals. His contemporary Maupertuis recognized the familial inheritance of human traits such as polydactyly and proposed that such heritable variations, arising through changes in germinal "corpuscles", might explain the diversity of species. There the matter rested, essentially until the rediscovery of Mendel's work, though the theme of microscopic structures carrying biological information was recurrent, as in Darwin's "gemmules" and Weismann's "biophors". Then within the space of two decades the correlation of Mendelian ratios with the behavior of chromosomes, and the demonstration of and mapping of mutations in Drosophila by Morgan and his school, completed the structure of classical genetics.

It was in the study of microorganisms that the application of genetic concepts at the molecular level began in earnest. There are two reasons for this: the nutritional requirements of prokaryotes are simple, and their large populations permit the observation of statistically rare events. The experimental ingenuity of Pasteur, and of other early giants such as Lister and Koch established an empirical tradition in microbiology which culiminated in the work of Beadle and Tatum on nutritional mutants. The treatment of populations of microorganisms as a kind of "biological microscope" for genetic purposes found further application in the work of Avery, Delbruck, Luria, and Hershey, which implicated DNA as the genetic material.

Genotype and phenotype.

In general the genetic structure, or genotype of an organism is simpler (more discrete) than the phenotype (which can sometimes present an apparently continuous range of variation). The use of cloned cell cultures for genetic analysis and to amplify molecular events undetectable at the level of the individual organism has been indispensable for observation of the genetic machinery and its products. One method has been to define biochemical phenotypes, to identify and map the relevant genetic loci, and to produce mutations at these loci. The mutants can then be used

to "change the hand that is dealt" -- i.e. to observe phenotypic permutations of the system under study due to quantal changes in its genetic information. In a sense the phenotype can be said to be mapped onto the genotype, but the combinatorial possibilities involving activation of different segments of genetic information and environmental effects confer both increased complexity and a degree of flexibility on the phenotype. This partial uncoupling of genotype and phenotype becomes more prominent in eukaryotic, multicellular organisms and raises specific questions in relation to phenomena such as development, differentiation, regeneration, and metamorphosis.

II. THE LOGIC OF THE CELL

Mendel's legacy: quantization and quanta in biology.

In some regards the figure of Mendel is more central with respect to the development of molecular biology than that of Darwin. In a view of life processes as consisting of molecular interactions and transformations the resolving power of genetics was of decisive importance. Molecular biology is really the child of genetics, and it is genetics which has quantized the conceptual elements of the science and permitted new theoretical insights. Conceptual quantization in modern biology begins with genes, their RNA and protein products, and the reactions catalyzed by the enzymes. Larger conceptual units may be regarded as built up of these primitive units: e.g. DNA replication complexes; operons with associated RNA polymerase, sigma factor, repressor, and cAMP-binding proteins; ribosomes, messenger RNA, and associated protein synthesis factors; multi-enzyme systems; and, ultimately, membrane systems, chromosomes and various cellular organelles. What is striking in all of these structures is that they represent functional as well as spatial/topographical units. A virus particle is quite clearly also such a unit. In research, conceptual quantization leads to a one-to-one comparison of genes, macromolecules and molecular transformations. These represent not only three types of experimental subject matter, but also three, mutually reinforcing, kinds of logic in experimentation: genetics, structural biochemistry and enzymology.

Molecular genetics developed from a complex cross-fertilization between microbiology and biochemistry, and in some ways is a distinctively American hybrid. The loci were East Coast and West Coast: first Morgan and then Avery in New York, then Morgan, and Beadle and Tatum in Pasadena, and after the war, the phage school at Cold Spring Harbor and at Pasadena. The exploitation of large microbial populations permitted a new level of sophistication in the selection of mutants of pre-defined phenotype, the

determination of precise linkage relationships, and construction of specific genotypes for research applications. Recombinational techniques, and new approaches such as transduction, sexduction and the use of conditional-lethal mutants have been continuously refined. The earlier, rather crude DNA transformation methods have been followed by newer, more powerful methods for direct infection of prokaryotes and eukaryotes with defined segments of DNA. If all of these developments had occurred as an extension of classical genetics (as in Benzer's fine-structure mapping of the rII region in T4), an exceedingly baroque, yet intellectually satisfying formal system might have arisen, but the connection of genotype and phenotype would have remained a matter for formal deduction. It was the interplay of genetics with the analysis of macromolecular structure and the enzymology of biosynthetic pathways that led to the integrated view of gene action which is now available.

Proteins were the first macromolecules to be deciphered. The colloidal theory of protoplasm which dominated at the beginning of this century was rudely shattered by the discovery that enzymes could be crystallized and that they have precise molecular weights (Sumner, Svedberg). Concurrently there developed at Cambridge the application of X-ray diffraction to the probing of macromolecular structure under Bragg and his colleagues. The history of this group and the unravelling of protein and nucleic acid 3-dimensional structures has already been told in some detail (Olby, Judson). Chemical approaches to the subunit bonding and sequences of these molecules, pursued at Cambridge (Todd, Sanger) and at the Rockefeller University, lent support to the X-ray method. Sanger's later work on nucleic acid sequencing together with methodologies developed at Harvard by Gilbert has permitted direct sequence comparisons of DNA and messenger RNAs with proteins.

The cradle of modern enzymology was in Germany. The school which grew up there was based on the strong tradition in the organic chemistry of natural substances in that country (Liebig, Fischer, Willstätter). A series of investigations (Warburg, Meyerhof, Krebs, and Lipmann) led to the isolation of pivotal enzymes and cofactors in the cellular glycolytic and respiratory apparatus and the demonstration that this apparatus is widely shared among different sorts of cells. Prior to World War II the emigration of many scientists led to transplantation of the main thrust of this work to the United States. This development coincided with the initial use of isotopic tracers at Columbia University, where the biochemistry department became heavily populated with emigrés, owing to skillful recruitment by Hans Clark (a pupil of Fischer). The post-war introduction of radioactive isotopes of carbon, phosphorus and hydrogen played a major role not only in the assay of enzyme reactions, but in defining the phenotypes of various mutants,and in the isolation of minute amounts of biosynthetic intermediates. The investment of large resources in biochemical research in the United States beginning in the late 1950's led to new advances. The ground had been prepared by the early involvement of Zamencnik and Lipmann in protein synthesis in vitro. The role of in vitro systems in dissecting the mechanism of protein synthesis and the genetic code is well known.

Later developments have included the detailed analysis of DNA replication and transcription systems, the cell-free synthesis of active enzymes, and the function of regulatory molecules *in vitro*.

The 3-fold heuristic of molecular biology

As set forth above, molecular biology has developed from the intersection of three experimental disciplines. It is usually a combination of these three approaches which has been responsible for nailing down an accepted model, and for demonstrating the relationship of a new phenomenon to the overall biology of the cell. A good illustration is transcription from bacterial operons. The mechanism was deduced from genetic discoveries at the Pasteur Institute, the isolation of radiolabelled messenger RNA species by groups at Cambridge, Cal Tech, and Harvard, and ultimately by the construction of *in vitro* systems for the demonstration of RNA synthesis (the discovery of RNA polynerase was almost simultaneous with the initial evidence for messenger RNA *in vivo*). Another, related example is the discovery of intervening DNA sequences and mRNA splicing, which grew out of the recognition of "silent" mutations in the spliced-out regions, the identification of RNA species with and without sequences corresponding to these regions, and the matching of these RNAs to defined DNA segments. In this case, the enzymology of the splicing event(s) is still awaited. Recent research in these areas has been greatly aided by the use of a new generation of macromolecular reagents, purified from cells, or synthesized *in vitro*. Restriction enzymes for analysis of DNA structure, defined nucleic acid sequences for hybridization experiments, and -- for other applications -- antibodies, including monoclonal reagents.

Genetic analysis, ultimately resulting in knowledge of DNA sequences, stands in a special relationship to the other, more purely biochemical approaches to the analysis of phenotype. The genotype specifies the phenotype, or at least sets limits within which the phenotype may vary. A DNA sequence as such, however, is a sterile object; it is only by interaction with the phenotypic apparatus that the information in the sequence gains expression. Thus in decoding the genetic message the cell must act as interpreter for the biologist.

III. FROM CELL TO ORGANISM

In studies of development and differentiation the heuristic of molecular biology encounters serious difficulties. No one-to-one correspondence exists between phenotypes such as morphogenesis or differentiated cell behavior and the presence or absence of a particular functional gene or protein.

Pleiotropism reigns in the genetics of cell-cell interactions and tissue construction.

Development

Some examples in the genetic analysis of development illustrate the complexity of the issues. Genetic determination in Drosophila metamorphsis has been intensively investigated during this century (e.g. Hadorn). Normal and abnormal patterns of imaginal disk determination and cross-determination reveal a complex hierarchy of genetic elements controlling emergence of adult tissues. The correlation of these genetic characters with specific protein products is a formidable task, but will undoubtedly be accomplished at some future date; the question which must be raised, however, is whether knowledge of these basic gene/protein correlations will be the most direct route to understanding how the system works.

The T-locus in mice has appeared to offer over the past several decades an almost perfect system in which to wed the genetic dissection of development with biochemical techniques. The T-locus is located on the same chromosome as the major histocompatability complex (MHC) "supergene" locus and has some similarities to this locus in representing a family of genes with related effects. In fact, the protein product of one of the mutant genes in this family is a cell-surface protein bearing a resemblance to some MHC products. It is of interest, however, that the initially promising serological analysis of t-mutants has failed to lead to a hoped for one-to-one relationship between protein products of linked genetic regions within the T-locus and defined steps in embryogenesis.

The early euphoria generated by the prokaryotic triumphs of molecular genetics has led to serious efforts to probe behavioral characteristics of metazoa by genetic methods. Noteworthy in this respect are the studies of Drosophila by Benzer, of the nematode nervous system by Brenner, and of neurological mutants of mice by Sidman's group. In each case it has proved impossible to separate morphogenetic aspects of nervous system development from behavioral effects as such. This is well illustrated by the work from Sidman's laboratory, where mutant mice with defects in locomotion have been shown to have gross derangements of cerebellar organization.

Differentiation

Hemopoiesis in the mouse is one of the most extensively studied differentiation systems, offering the possibility of experimental attack by immunological, biochemical and genetic methods. The complex role of the MHC in genetic regulation of humoral and cell-mediated immune responses through the protein products of sub-regions within the MHC (H-2 proteins and Ia proteins) has been established by the use of genetic variants and recombinant and congenic mouse strains. Here is a system which specifies not only whether or not a mouse can make antibody against a particular foreign

antigen, but also whether various differentiated T- and B-cell types can interact so as to modulate the immune response. Immunogenetics has defined stages of T- and B-cell differentiation in terms of cell-surface antigen profiles. The same methods have also led to the description of hormones active in transitions between stages. The theory of cellular commitment to a particular differentiation pathway (Holtzer) and the triggering of this commitment by extracellular effectors has proved fruitful for this work. In the case of B-cells the concept of commitment has been provided an ironclad foundation with the demonstration of DNA sequence translocations for immunoglobulin genes. Recently Boyse and colleagues have found a genetic linkage of mating preferences in mice to the MHC locus. Thus a central nervous system response can also be directly connected with the pleiotropic effects of the MHC.

A beautiful cameo example of pleiotropy in hemopoiesis is provided by the W and Steel mutants of mice. Here two unlinked loci have been found to affect erythroid cell maturation, coat color, and spermatogenesis by what appear to be two mechanisms: the phenotype of a precursor cell (to erythrocytes, melanocytes, or male germ-line cells) or the phenotype of a tissue milieu which promotes (or in the case of mutants, prevents) differentiation. This type of pleiotropy holds out the promise of a single genetic program uniting the development of various tissue types. (There are other genetic data suggestive of a connection between hair follicle development and T-cell development).

Relatively high frequencies of mutation have been observed for the T-locus, the MHC, the W and Steel system, and (in in vitro selective systems) for immunoglobulin genes. It is possible that DNA rearrangements, which play a role for immunoglobulin synthesis, may have a wider relevance for generation of abnormal rates of mutation.

Cancer

Cancer is a disease of development resulting from a somatic genetic change. The necessity for a genetic alteration is documented by overwhelming evidence, from the stability of the neoplastic phenotype to the specific interaction of oncogenic agents (viruses, chemicals) with the genetic apparatus of cells. The master genes involved in transformation events appear to be multiple. This can be deduced from the ratio of transformation to mutation rates in cultured cells, and from the emerging evidence that oncoviruses can incorporate a number of different host genes into viral genomes, creating thereby agents with different spectra of transforming capacity for differentiated tissue types. Thus various leukemia viruses can transform cells of the T-cell lineage, or B-cell precursors, or cells of the erythroid, the macrophage, or the granulocytic pathways of differentiation. Other, related viruses attack mesodermal cells, producing sarcomas. Tissue-specificities are not absolute, but the suggestion of virus-encoded, host-derived genetic information interactive with specific tissue types is tantalizing, and a subject of intense current investigation.

The balance between growth as a relatively undifferentiated precursor cell and continuous differentiation into terminally differentiated, non-neoplastic cells of the same lineage is a characteristic of many cancers. A dramatic example is afforded by the ability of mouse embryonal carcinoma cells to differentiate normally in the environment of a mouse blastocyst and contribute to various tissues of a normal adult animal. Cancers of the blood-forming system may produce wantonly growing cells of one lineage (e.g. lymphoid) and, as determined by presence of a marker chromosome, contribute to the supply of normal cells of a different lineage (e.g. red blood cells). Friend leukemia cells from the mouse, induced to differentiate and produce hemoglobin, cease dividing. Solid tumors of the breast or colon in the human can be shown to throw off differentiated cells typical of the organ involved, which are no longer capable of sustained growth. Besides offering an entry point for the design of new therapies aimed at causing tumor cells to differentiate en masse, these observations underscore the complex role of genetic changes causing cancer in relation to the activity of the cell genome in carrying out particular differentiation programs.

IV. GENETICS AND BIOLOGICAL EXPLANATION

Genes in evolution.

In a very precise sense genes are the core components of evolution; since genes are what survive from generation to generation. Darwin's rule requires first of all the natural selection of genes. To this must be added, however, that genes scarcely ever survive (except in the case of some defective viral genomes) as isolated units; they are clustered in the species genome, and survival is at the level of the individual of the species. Genes mutate singly as a rule, and their persistence can be measured in the gene pool of a species population, but an adaptive trait also increases the survival probability of the associated genetic traits which collectively define the species. Thus complex phenotypes such as sexual mating do not simply enhance selection of single genes, despite the frequency of cross-over events, but increase the survival capacity of the genetic framework (species genome) within which individual genes mutate and individual organisms live and die. (I use the term "species genome" in the same sense that a species phenotype can be recognized as a basis for classification.) Altruistic behavior is not merely clannish, in benefiting the genotypes of siblings and first cousins, it operates in the interests of the species genome. Mechanisms for rearrangement of DNA sequences and selection of mutator genes, permitting rapid changes in phenotype or emergence of new species also have survival value for gene sets.

Just as enzymatic conversions present the dynamic aspect of the generality of chemical structures which are isolable from cells, evolution provides a time dimension to the set of DNA base sequences which collectively define the

species genome at a given point in time. " Time's arrow" is usually taken to mean the entropic drift of Second law dominated processes. It can also be applied to the temporal sequence of DNA structures as these have arisen in evolution. The Second Law and the Principle of Natural Selection are, for the purposes of the biologist, logically complementary. In its statistical form the Second Law states that any (closed) system will spend most time in those macro-states for which the ensemble of energetically equivalent micro-states is the largest in number -- that is, the system will move, in time, toward the most probable states, which happen (because of the dense configuration of micro-state-space for such an outcome) to be what we would see as disorderly or random ones. No one who has seen a teenager's room need doubt the validity of the Second Law. Natural Selection states that evolution moves toward more complex - orderly- forms because of selection of those genotypes which are the fittest to survive.* Predominancy of the "most probable" states and the "fittest" life forms might appear as tautological propositions but both are synthetic and emerge from the physical structure of matter. Thermodynamics would be sterile without the existence of real physical systems. Life with its genetic mechanism manifestly exists; it has sprung from matter itself, and its space-time distributions, as inexorably as the birth and death of stars.

Quasi-genes

Although in one sense the organism is summarized in the genome (hence the fascination of "cloning" higher organisms), yet DNA is mute without cell machinery to decode it. Cells can only derive from other cells (Virchow). The interaction of chromosomes with the cells which they inform is a dynamic one: the double helix "breathes" to allow transcription; chromatin unfolds and puffs; a host of molecular emissaries come and go. The effectiveness of genotype translation into phenotype, the very survival of the organism, depend on an exquisitely modulated, flexible interaction, not a master-slave operation. Information in DNA must be compatible with the cell functioning in a variety of different ways, depending on environmental signals. Environment can curb even the nucleus of a malignant cell (e.g. for embryonal carcinoma cells, cf. above), as well as shape cells into differentiated tissues, induce specific enzymes, and maximize energy utilization. The most successful genetic program will permit the organism to function in a large number of modes. In higher organisms the development of sense organs, locomotion, and a central nervous system permit a much wider range of adaptation to particular situations, within the same genetic program.

---

* From the thermodynamic point of view living systems succeed in limiting entropy at temperatures permissive to considerable rates of (enzyme-catalyzed) reactions. As Schroedinger pointed out, the result is a narrowing of permissible states, analogous to what is achieved for a perfect gas by subtracting energy (lowering temperature)in a system. The energy tied up in a living cell is not very large; more energy has been spent in exothermic reactions to build the organization up. Externally derived energy (ultimately solar), genetic flexibility and ingenuity, and the cooperative properties of matter have permitted the evolution of ambient-temperature systems with the functional precision of zero-degree machines.

The emergence of higher organisms has resulted in new systems of information storage and retrieval, which can be termed "quasi-genetic". The information in these systems is not represented in DNA, yet has permanence and capacity to be propagated. The brain is the most obvious such system, but other functions such as the immune response invite comparison. Though ultimately dependent on DNA expression, memory and learned behavior of an individual probably inhere in a different sort of physical structure, which can also be "translated" into manifest functioning. The empinging of environmental signals, the requirement for finely tuned and flexible responses, are similar to what are asked of the DNA apparatus. On another level social structures, whether of ants, wolves, or people, exhibit stability, information storage, and refinement of behavior leading to enhanced survival of the species. Though dependent in the first instance on the central nervous systems of individuals (and thus still ultimately on DNA) societies have also created information-laden structures (homes, habitats), and, in the case of humans, formal information systems (libraries, computers) which are external to biological organisms. Through the mediation of human organisms these systems are also duplicated, and by the translation of coded information into overt effects can radically alter the conditions of natural selection. In human societies the practice of science and the enjoyment of art depend on collective enterprises and the transmission of ideas by non-genetic routes, and these creations descended from human DNA have changed the circumstances for survival of the species and of its DNA. More simply put, the context (cytoplasm, organism, culture) in which DNA is expressed is crucial for the biological result.

The limits of explanation.

The bacterium E. Coli contains about 4000 genes. Human beings probably have at most 10 times this number of genes,and they may have less. Most of the additional genetic information in humans and other eukaryotic organisms is needed for the specification of specialized tissues and organs with (near) duplication of many genes involved in some systems, such as immune response. Before the advent of modern DNA cloning and sequencing methods and the application of computers to biological research, the prospect of completely unravelling the structure of a human genome seemed remote. Now such an achievement seems only decades away. When we possess knowledge of such DNA sequences, and of the gene products which they program (aided by more refined methods for in vitro transcription and translation) will this provide us with a description of a human being? It is likely that the available number of "slots" in the genetic tape will turn out to be too limited in number to accomodate many aspects of human behavior (much less ideas) as inherited traits. It will be easier to fit into a limited genetic repertoire some highly specific reflexes, as well as generalized pleasure/displeasure responses to threshold stimulation through specific sensory pathways. In other respects we may well turn out to have "wired" ourselves, using the ample supply of neurons at our disposal.

There are three ways in which the genetic constraints are not as tight as might at first appear, howwever. The first is due to the fact that gene interactions at the phenotypic level multiply possibilities, creating additional phenotypes. Thus, as already outlined, a single genotype will permit a large number of states of an organism, depending on environmental effects. Secondly, direct gene interactions at the DNA level have been observed: for example, a single constant region immunoglobulin gene can fuse with a number of variable region genes. Somatic rearrangements in DNA clearly "stretch" the available information by specifying gene products not originally coded. A third consideration is that interaction of the genotype with the environment educates and refines the responses of the organism. Watch a dog trained as a retriever chase a stick. The exquisite control of motor functions is partly inbred, but it is also due to incessant practice of certain movements, comparable to the musician perfecting his art. In many situations refinement of a performance (which may be programmed by a modest number of genes) is really the essence of what might appear as an incredibly intricate piece of behavior. Environmental effects may play a role in the much-publicized specialization of cerebral hemispheres in humans: in women, for whom cultural emphasis on aggressive manipulation of the environment (building, working with tools, etc.) has been less pronounced than for men, the usual differentiation of the left (right-handedness) hemisphere is less marked. It seems likely that some types of behavior which have been singled out by sociobiologists as candidates for genetic transmission may turn out to reflect the interaction of a simpler, or more diffuse "core" trait with factors extant during the lifetime of the individual.

Conclusion.

The quantized nature of biological systems and accumulating knowledge of the mechanisms by which macromolecules specify complex phenotypes, raises the possibility that before very long we shall know most of the essential facts about how cells function and cooperate in the development of multicellular organisms. This possibility has been discussed by Stent in terms of a kind of twilight of molecular biology -- the completion of the science, with certain problems, such as the molecular correlates of conscious thought processes, inaccessible. We have no equivalent of a von Neumann model for an organism completely understanding itself, certainly not at the moment of action. If it turns out, as appears likely, that we ourselves, and the human society to which we belong, determine the nature of what we think, through "self-wiring" based on our actions and responses to experience throughout our own lifetimes, then we are the protagonists of the drama, and we cannot evade the existential choices. In accepting this responsibility we assume our true roles as bearers, and in a real sense, creators of the evolutionary process.